Module 11 - Probabilities and Statistics

ME3023 - Measurements in Mechanical Systems

Mechanical Engineering

Tennessee Technological University

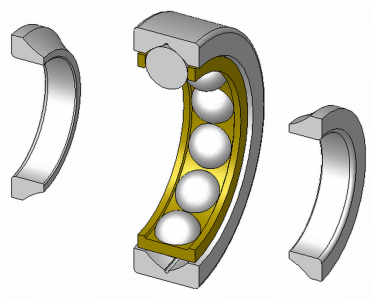
Topic 1 - Histograms and Probability Density Functions

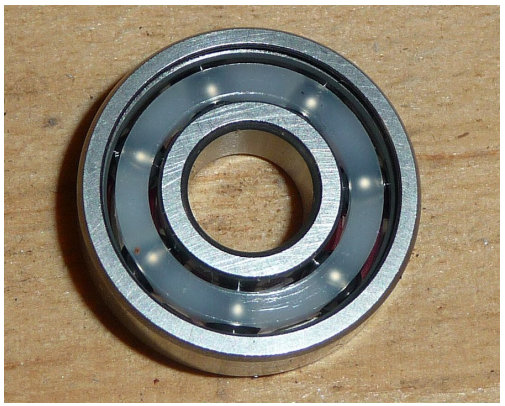
Topic 1 - Histograms and Probability Density Functions

- A Population of Data
- Randomly Distributed Data
- Frequency Histogram
- Probability Density Function

A Population of Data

Consider a manufacturer that makes 100000 of ball bearings in a set.





- How does the manufacturer ensure the quality of the product?
- How the does the seller communicate this to the buyer?
- How does the engineer use this information?

Images: Wikipedia

A Population of Data

The data set, or population, is generated by taking measurements of individuals chosen randomly from the entire population.

Sampling refers to repeated measurements of the **measured variable** under fixed operating condition.

$$X = \{x_1, x_2, x_3, \ldots, x_i, \ldots, x_{n-1}, x_n\}$$

Randomly Distributed Data

Our discussion will assume that the values in the data set are **randomly distributed** about a mean value. It is important to consider what this means.

Frequency Histogram

"A histogram is an accurate representation of the distribution of numerical data. It is an estimate of the probability distribution of a continuous variable (CORAL) and was first introduced by Karl Pearson.[1] It differs from a bar graph, in the sense that a bar graph relates two variables, but a histogram relates only one. " - Wikipedia

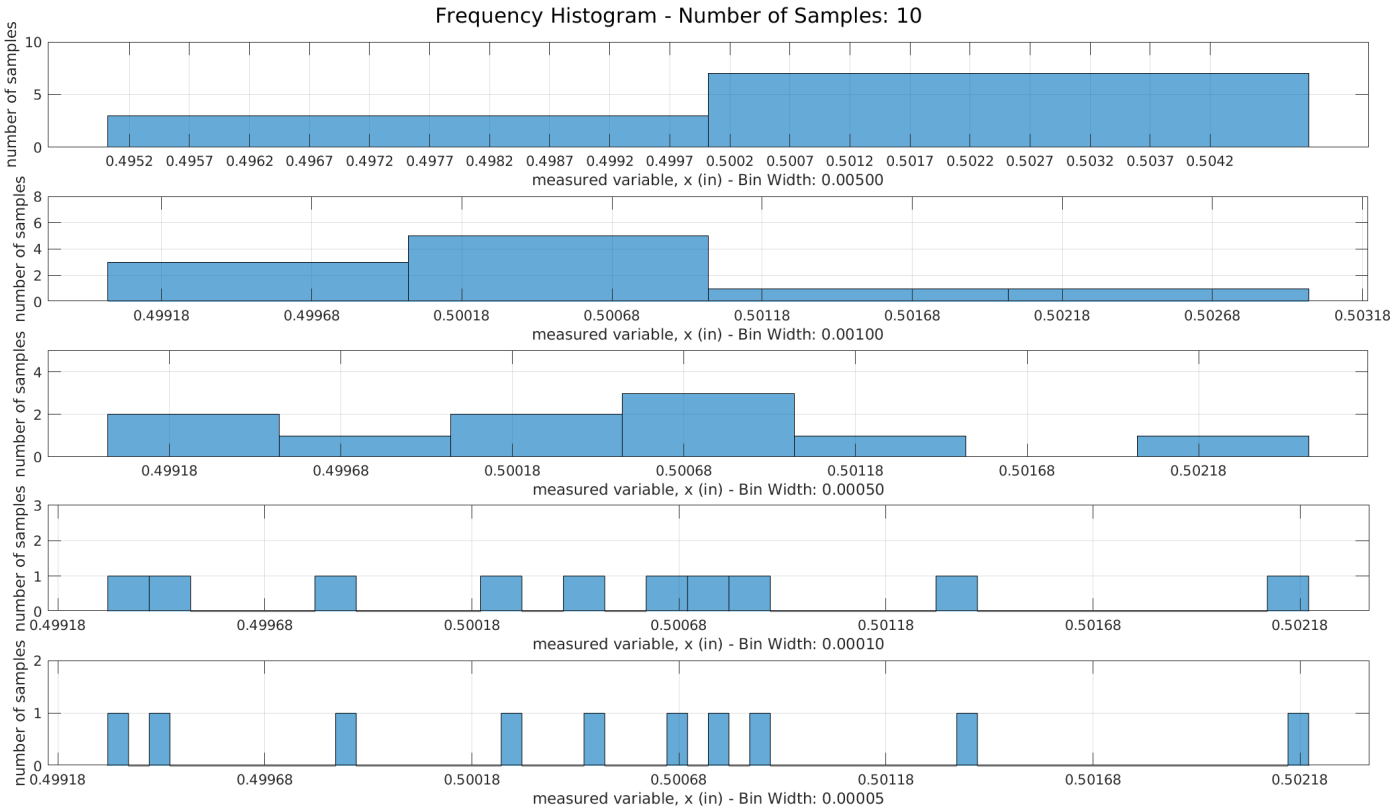
Frequency Histogram

Consider sampling the 0.5 inch diameter ball bearings.

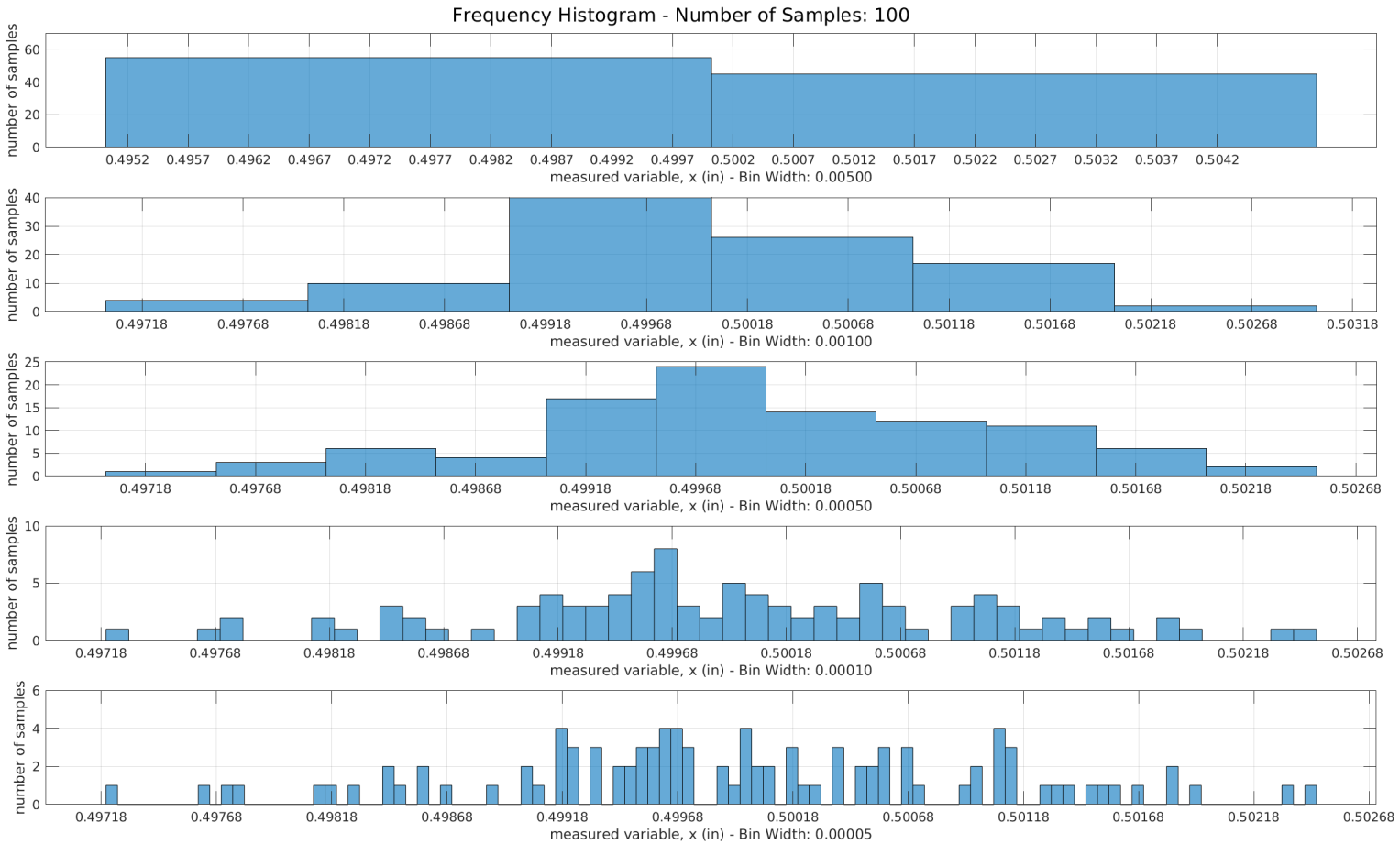
0.5016	0.4991	0.4981	0.5003	0.4988	0.5007	0.4994	0.5000	0.4999	0.4999	0.4988	0.5007	0.4992
0.4997	0.4991	0.4986	0.5000	0.5018	0.4996	0.5010	0.4991	0.5011	0.5006	0.5006	0.4996	0.4997
0.5014	0.5001	0.4997	0.4994	0.5001	0.5009	0.5009	0.4992	0.4983	0.5004	0.5005	0.5011	0.5010
0.5004	0.4992	0.5009	0.5003	0.5018	0.5004	0.4996	0.5014	0.4992	0.4986	0.5002	0.4981	0.4999
0.4986	0.4992	0.5003	0.5002	0.4999	0.4998	0.5013	0.5009	0.5003	0.4986	0.4991	0.4997	0.5004
0.5030	0.4995	0.5012	0.5005	0.4992	0.4995	0.5005	0.5000	0.5005	0.4993	0.4987	0.5006	0.4997
0.4992	0.5018	0.4976	0.5003	0.4991	0.5007	0.5002	0.4980	0.5004	0.5000	0.4997	0.5005	0.4997
0.5002	0.5004	0.4980	0.5011	0.5003	0.5013	0.5000	0.5001	0.5003	0.5000	0.5013	0.5000	0.4997
0.5004	0.4988	0.5003	0.4999	0.5004	0.4996	0.4999	0.5000	0.4993	0.5003	0.4996	0.5012	0.5005
0.4992	0.5008	0.4997	0.5004	0.4998	0.4998	0.5006	0.5002	0.5010	0.5014	0.5019	0.5013	0.4998
0.5008	0.4997	0.5002	0.4993	0.4995	0.5007	0.4986	0.5009	0.4999	0.5002	0.5010	0.4993	0.5004
0.4996	0.4993	0.4982	0.4994	0.4993	0.4995	0.5002	0.5004	0.4999	0.5005	0.4993	0.5003	0.5008

Data: Generated in MATLAB, see probabilty_statistics_topic1.m

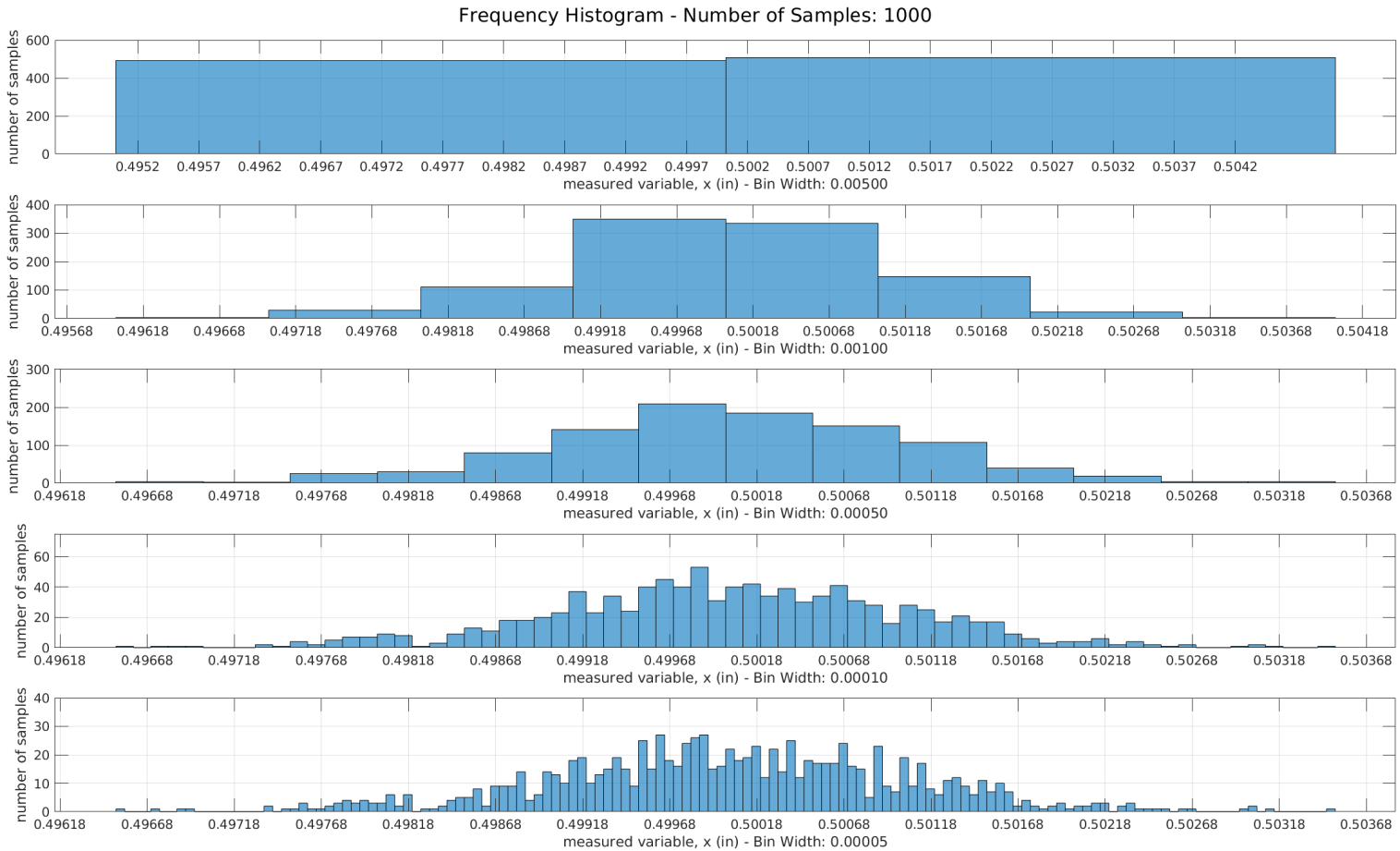
Frequency Histogram



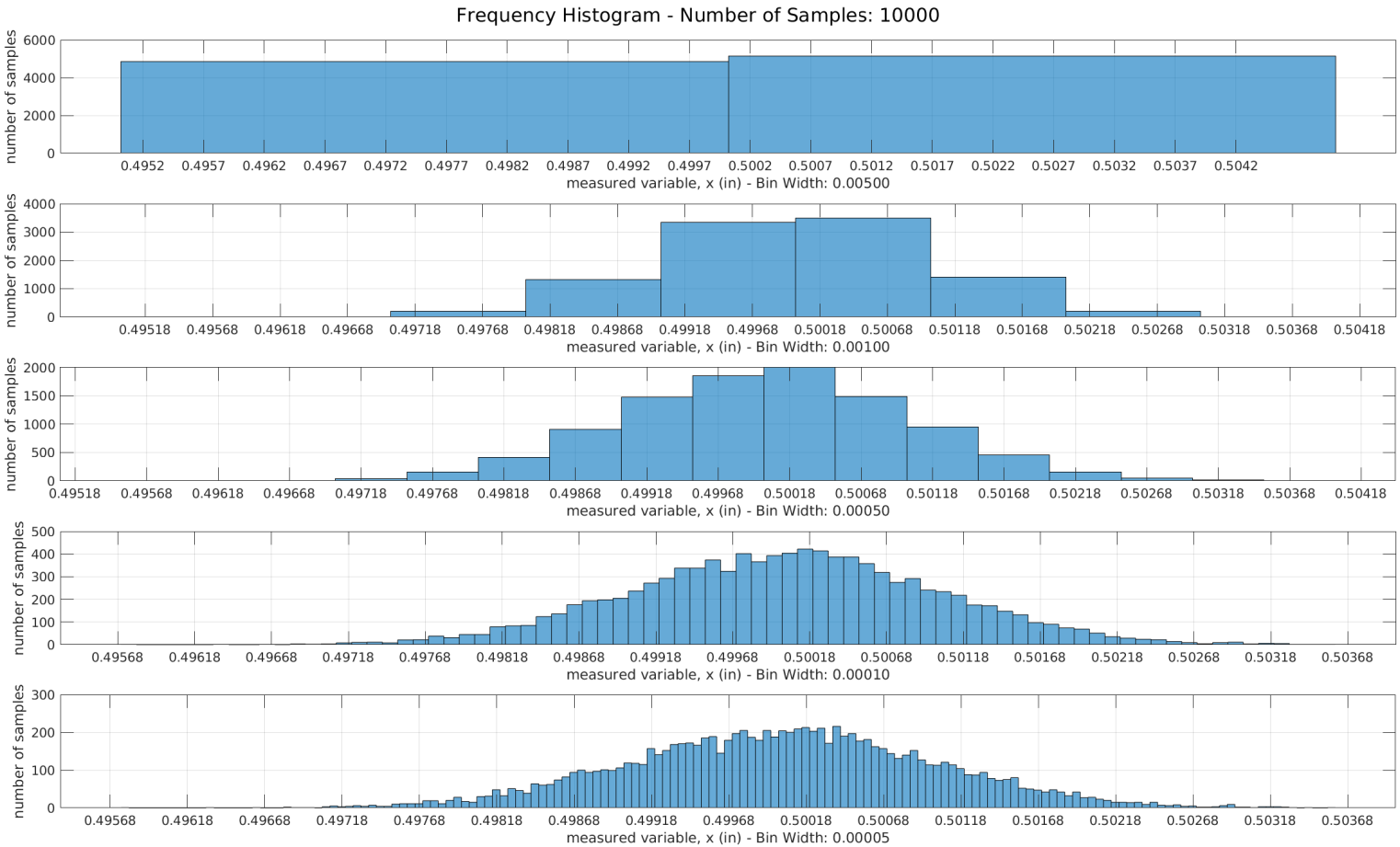
Frequency Histogram



Frequency Histogram



Frequency Histogram



Probability Density Function

"... a probability density function (PDF), or density of a continuous random variable, is a function whose value at any given sample (or point) in the sample space (the set of possible values taken by the random variable) can be interpreted as providing a relative likelihood that the value of the random variable would equal that sample ... the PDF is used to specify the probability of the random variable falling within a particular range of values, as opposed to taking on any one value. This probability is given by the integral of this variable's PDF over that range—that is, it is given by the area under the density function but above the horizontal axis and between the lowest and greatest values of the range ..." - wikipedia

Probability Density Function

- The frequency with which the measured variable assumes a particular value or interval of values is described by its **probability density function**.
- If a **central tendency** exists we should be able to see this in the **probability density function**.
- As binsize of the **histogram** of the data set goes to zero this becomes the **probability density function**.

Probability Density Function

Now, lets's do an example.